Machine Learning, Statistiques et Programmation

Version 0.1.407

Xavier Dupré

oct. 27, 2018

Table des matières

CHAPITRE 1

Introduction

J'ai commencé ce site un jour de nuit blanche, après avoir codé toute la journée et échangé avec mes collègues via des codes review. Les mails qui tombent sur mon portable, le énième commentaire sur la position des espaces dans mon code, une énorme angoisse m'étreignit alors. Les lignes de code défilaient comme les blocs de Tétris dans ma tête. Je me revois alors après une nuit blanche sur ce jeu, plus excité encore qu'après avoir du café à la fontaine. Je me suis dit qu'il fallait que j'arrête et que je remette un peu de sens dans tout ça.

Derrière le code, il y a les algorithmes et derrière encore des idées. Plein d'idées magnifiques, plein d'intuitions. Un code efficace n'a rien à voir avec sa lisibilité. C'est même souvent le contraire et il ne sera jamais lisible. Mais l'idée soujacente, elle, lorsqu'on l'a comprise, elle devient très clair. On n'a même plus besoin de l'écrire.

Il faudra probablement quelques années avant que ce site ne devienne conséquent, voire exhaustif. Il faut bien commencer quelque part. J'aurais pu écrire des pages wikipédia mais je n'aurais pas pu y mettre du code, des notebooks. J'ai commencé ce travail en latex lors de ma thèse, je me suis aperçu qu'il me restait quelques démonstrations à terminer.

Et en français... J'enseigne en français et puis l'offre est tellement riche en anglais.

Installation

Les exemples de codes sont disponibles sous la forme d'un module python et des *notebooks* (page 299) accessibles sur le site.

```
pip install mlstatpy
```

Je n'ai pas la prétention d'être exhaustif. Vous la trouverez dans la documentation des librairies de machine learning qui implémentent la plupart des algorithmes performants connus. Allez voir scikit-learn [1] pour connaître ce qui marche. La documentation rend obsolète n'importe quel livre au bout de six mois.

1. http://scikit-learn.org/stable/

Ceci est un exemple, cliquez sur le lien de téléchargement pour obtenir le cours complet.